

CONSUMER TECHNOLOGY

# Why Twitter Can't Shake Its Harassment Problem

The most public-facing social network struggles to find a balance between free speech and user safety

By **NATHAN OLIVAREZ-GILES**

Aug. 24, 2016 9:27 a.m. ET

Twitter Inc. has been working hard lately to shake its reputation as a haven for harassment.

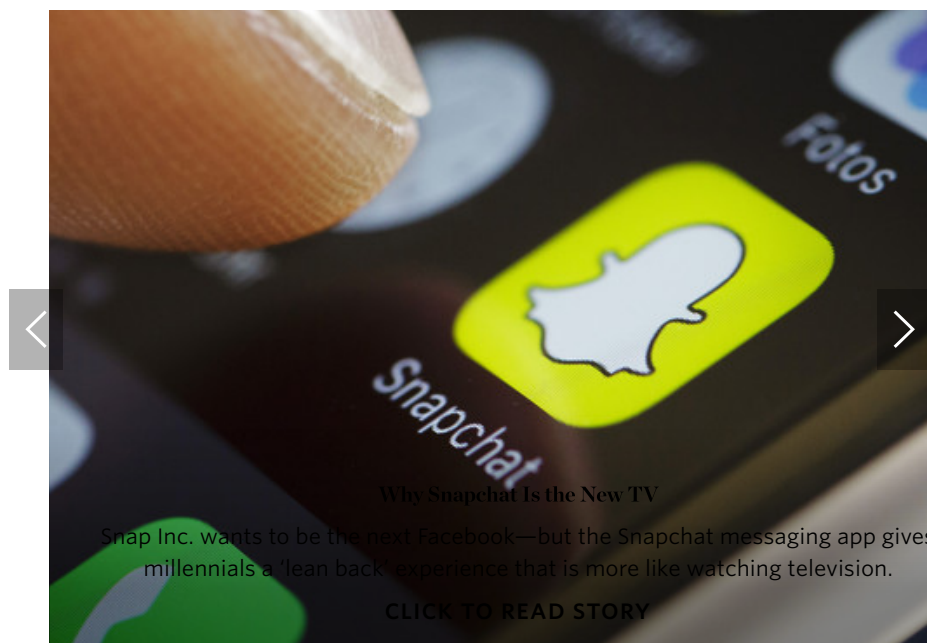
The social network recently banned a notorious troublemaker and instituted a “quality filter” to limit responses users can see. Earlier this year, it improved its abuse reporting system and convened a Trust & Safety Council of outside advisers. A Twitter spokeswoman told the Journal that the company has “invested heavily in improving our tools and enforcement systems” to respond to incidents of abuse and prevent future ones.

But the increase in effort may have little effect, experts say. The problem is so widespread, Twitter’s own design can publicize and magnify instances of harassment, and the company’s rules and response tactics are still too vague and inconsistent.

For many users, Twitter has improved, said Paul Booth, an associate professor at DePaul University who researches how people use social media. “There are fewer trolls and fewer hate groups on there, and abusive people can get away with a lot less.”

---

## MORE FROM TECH



Still, “if you’re a woman, or if you’re a celebrity of color, Twitter is still full of vitriol and can be a place where you run into a lot of hate directed your way,” Mr. Booth said. “For those people, Twitter doesn’t feel safer.”

In July, a conservative blogger named Milo Yiannopoulos targeted a series of harassing tweets at “Ghostbusters” actress Leslie Jones while encouraging dozens of other Twitter members to do the same. The traumatized Ms. Jones shared some of the sexist and racist tweets she received, then quit Twitter temporarily. The social network banned Mr. Yiannopoulos for life and deleted his tweets. While harassment stories like this aren’t uncommon, Twitter’s use of a lifetime ban on such a high-profile Twitter user was.

Twitter’s inherent design contributes to the problem.

“The abuse that happens on Facebook and Instagram and other places is hidden behind logins. It’s not viewable to the public and searchable online the way Twitter is,” said Karen North, the director of the Annenberg social media program at the University of Southern California. The public nature of Twitter makes harassment more visible and biting, she said.

But while Twitter’s openness makes people think of it as a public utility and a place for free speech, Ms. North said the company needs to remember it’s a private business.

“Twitter can ask people to leave, just as a restaurant and any other business has the right to refuse service,” she said. “It’s getting better at this, but it needs to be more aggressive at kicking out people who break its rules.”

Mr. Booth argued the rules themselves aren’t clear enough. “What counts as hate speech? What counts as defamation? What counts as trolling? We don’t have definitive definitions for any of these things in society, much less on Twitter,” he said. “How do you know when someone has crossed the line if you don’t know where the line is drawn?” (Twitter actually bans harassment and “hateful conduct,” which concern inciting abuse or promoting violence. It doesn’t specifically ban hate speech itself. Here’s its most recent rules of conduct.)

“We are reviewing our policies to prohibit additional types of abusive behavior and allow more types of reporting, with the goal of reducing the burden on the person being targeted,” the Twitter spokeswoman said.

Facebook Inc. has taken steps to curb abuse on its namesake social network as well as its photocentric subsidiary, Instagram. In July, Facebook’s global head of safety admitted that the company wasn’t doing enough to combat harassment and that it’s examining whether current tools and policies were effective. Instagram bans hashtags related to racism and body shaming, and in July, it gave users the ability to turn off photo comments.

Anne Collier, executive director of the Net Safety Collaborative anti-cyberbullying nonprofit and a member of Twitter’s Trust & Safety Council, said Twitter continues to search for a balance between protecting users and enabling free speech. No social network—not Facebook, not Instagram, not Snapchat—has found the perfect mix, Ms. Collier argued.

“I can’t emphasize enough how hard a problem this is,” she said. “If you’re a feminist activist and you get harassed all the time and Twitter responds by improving blocking and you can block a bunch of your harassers en masse. That is great, right?”

But there’s a downside, too, Ms. Collier said. “If you block somebody on Twitter, you no longer see what they’re saying about you,” she said. “Sometimes you need to see what others are saying about you so you can protect yourself offline.”

User safety will always be a work in progress at Twitter, Ms. Collier said, because the nature of harassment is that it’s unique to the people involved and the context of the moment. “This comes down to how people treat one another,” she said.

**Write to** Nathan Olivarez-Giles at [Nathan.Olivarez-giles@wsj.com](mailto:Nathan.Olivarez-giles@wsj.com)

Copyright ©2017 Dow Jones & Company, Inc. All Rights Reserved

This copy is for your personal, non-commercial use only. To order presentation-ready copies for distribution to your colleagues, clients or customers visit <http://www.djreprints.com>.